UK International Development
Partnership | Progress | Prosperity

I-SAPS Institute of Social and Policy Sciences
*Informing Policies, Reforming Practices*

# AI-ENABLED ADAPTIVE LEARNING

## A pilot for Elementary Schools in Punjab

# AI-Enabled Adaptive Learning

## A pilot for Elementary Schools in Punjab

***Bridging Technical Assistance for Governments (B-TAG)***

## Table of Contents

**List of Figures:**

# 1. Executive Summary

Punjab continues to face learning gaps among elementary school students despite various interventions. Traditional classroom teaching struggles to address wide differences in student learning levels. To address this challenge, Bridging Technical Assistance for Governments (B-TAG) programme, provided technical assistance to the Punjab School Education Department (SED) and the Punjab Education Curriculum, Training and Assessment Authority (PECTAA) to design, develop, and pilot an Artificial Intelligence (AI)-enabled adaptive learning solution for elementary schools in Punjab. The pilot was implemented in 50 public schools in the Sahiwal district, and targeted Grade 8 students in the subjects of English, Mathematics, and Science.

The intervention sought to improve Student Learning Outcomes (SLO) through personalised instruction. Using students' baseline assessment results, the platform generated adaptive learning pathways tailored to individual learning needs and continuously tracked student progress. The pilot combined technology-enabled learning with structured teacher support to ensure effective implementation within existing school systems.

This report documents the design, implementation, and evaluation of the pilot. Key components included the development of curriculum-aligned assessment items linked to prioritised SLOs, deployment of a web-based AI learning platform, and capacity-building support for teachers. Assessment instruments and learning content underwent rigorous review and validation processes by PECTAA to ensure alignment with the Punjab curriculum and quality standards.

Participating schools were selected using predefined eligibility criteria ensuring hardware availability and connectivity. Teachers received targeted training on platform utilisation, assessment administration, and strategies for supporting student engagement during scheduled learning sessions. Students completed baseline assessments prior to using the platform, while endline assessments were administered at the conclusion of the intervention period using standardised procedures.

The pilot was evaluated through a quasi-experimental design incorporating treatment and control groups, pre- and post-assessments, and complementary qualitative evidence from the stakeholders. The findings demonstrate that students who participated in the AI-enabled adaptive learning programme achieved consistently greater learning gains than their peers in control schools across all three subjects. Average scores in treatment schools increased by:

- 2.86 marks in English,
- 3.39 marks in Mathematics, and
- 3.04 marks in Science.

In comparison, students in control schools recorded gains of 1.60 marks in English, 1.22 marks in Mathematics, and 1.33 marks in Science.

Difference-in-Differences (DiD) analysis indicates that, after accounting for improvements observed in control schools, the intervention was associated with additional gains of 1.26 marks in English, 2.17 marks in Mathematics, and 1.71 marks in Science. Expressed as percentage-point improvements, treatment group students outperformed control group trends by 5.9 percentage points in English, 9.4 percentage points in Mathematics, and 7.3 percentage points in Science. Further analysis revealed that overall, boys performed better in Science and Math, while girls performed better in English.

Beyond learning gains, the pilot demonstrated the feasibility of integrating AI-enabled adaptive learning technologies within the public education system. The intervention highlighted the potential of personalised learning pathways, real-time progress monitoring, and data-driven instructional support to address diverse learning needs at scale.

The report concludes with key lessons emerging from implementation and provides recommendations for strengthening, refining, and scaling AI-enabled adaptive learning solutions within Punjab's education system. These recommendations focus on improving implementation effectiveness, enhancing teacher support mechanisms, strengthening technology infrastructure, and leveraging assessment and learning data to inform future education reforms.

**Scope and Reach**

District: Sahiwal

Intervention Schools: 50 (35 female, 15 male)

Teacher Trained: 78 (52 female, 26 male)

Grade 8 Students benefitted: 1,425 (1,064 girls, 361 boys)

Pilot Duration: Approximately 1 month

*Selected Control Group*

Schools: 50 (35 female, 15 male)

Grade 8 Students: 1,307 (935 girls, 372 boys)

# 2. Introduction

Despite significant progress in expanding access to education, public education systems in Pakistan continue to face persistent challenges in improving student learning outcomes. Evidence from national assessments and large-scale learning data, consistently indicates that many students progress through the education system without acquiring the expected competencies in core subjects. These learning gaps often reflect cumulative deficits from earlier grades, variations in instructional quality, overcrowded classrooms, and limited opportunities for teachers to provide differentiated support to students with diverse learning needs.

The selection of Grade 8 is very strategic from different standpoints. As the final year of elementary education, Grade 8 represents a critical transition point before students enter secondary education and begin studying more advanced subject content. Learning gaps that remain unaddressed at this stage often become increasingly difficult to remediate and may adversely affect students' performance in secondary school and subsequent board examinations. Grade 8 students are also likely to possess the digital skills necessary to engage independently with technology-enabled learning platforms, making them well suited to pilot adaptive learning approaches. From an implementation perspective, many high schools serving Grade 8 already have functional computer laboratories and basic internet connectivity, providing the minimum infrastructure required for digital learning.

However, classroom-level assessment practices often provide limited insight into individual learning needs, making it difficult for teachers to identify specific learning gaps and provide targeted support. Consequently, substantial variation in student learning levels within the same classroom frequently remains unaddressed, contributing to widening achievement gaps over time.

Globally, digital learning technologies and artificial intelligence (AI) have emerged as promising tools for supporting personalised learning and improving educational outcomes. AI-enabled learning systems can utilise diagnostic assessments, adaptive content delivery, and real-time performance data to tailor learning experiences to individual students. When integrated within formal schooling systems, such technologies can complement teacher-led instruction by providing actionable insights into student learning and enabling more targeted and responsive teaching practices. Recognising this potential, the B-TAG programme has been exploring the use of AI-enabled solutions to strengthen teaching and learning within public education systems.

In an earlier pilot, B-TAG supported PECTAA to implement AI-powered learning platform designed to assess student learning levels, monitor progress, and generate customised instructional recommendations for teachers in early grades. The platform's recommendation engine analysed student performance data and suggested learning activities aligned with individual needs, enabling teachers to make more informed pedagogical decisions and provide targeted support. The pilot began with diagnostic assessments administered through the platform, which generated detailed profiles of individual student learning levels. Based on these results, the system recommended tailored lesson plans and learning activities for teachers. Implemented across 30 schools in Lahore, the first phase of the intervention reached

1,268 Grade 1 students and resulted in an increase in the proportion of students achieving the highest learning level from 51.6% at baseline to 69% at endline. A second phase, involving 1,407 Grade 1 students, recorded a rise in the proportion of students at the highest learning level from 15.14% to 37.98%. These results provided initial evidence of the potential of AI-enabled personalised learning approaches to improve student outcomes within public school contexts.

Building on these experiences, B-TAG, in collaboration with the SED and PECTAA designed and piloted an AI-enabled adaptive learning intervention for elementary schools in Punjab. Unlike the earlier teacher-facing model, which generated instructional recommendations for the teachers, the new intervention was designed to provide students with personalised learning pathways directly through an adaptive digital platform. Based on students' performance in baseline assessments, the platform automatically adjusted learning content and activities to match individual learning needs while continuously tracking progress over time. The focus on Grade 8 reflected both the educational significance of this transition year and the practical feasibility of implementing technology-supported learning in schools with access to computer laboratories and basic digital infrastructure.

The rationale for the intervention was grounded in three interconnected considerations.

i) There was a need for more granular diagnostic information to identify individual learning gaps among middle-school students.
ii) Public schools required scalable approaches for providing differentiated learning support within classrooms characterised by diverse learning levels.
iii) There was a need to generate evidence on the feasibility, effectiveness, and operational requirements of implementing AI-enabled adaptive learning solutions within government school systems.

This report documents the design, implementation, and evaluation of the AI-enabled adaptive learning pilot. It presents key implementation processes, including intervention design, school selection, assessment and content development, teacher orientation and training, platform deployment, and the administration of baseline and endline assessments. The report also examines programme outcomes and implementation experiences to inform future efforts aimed at integrating adaptive learning technologies within public education systems in Punjab and beyond.

# 3. Methodology

## 3.1 Overall Evaluation and Documentation Approach

The evaluation framework adopts a quasi-experimental design incorporating treatment and control groups, with student learning measured through standardised assessments administered at baseline and endline. This design allows for comparison of changes over time between participating and non-participating schools. Quantitative assessment data are complemented by qualitative inputs from teachers and students, collected to contextualise observed patterns and to document experiences related to platform use, assessment administration, and classroom facilitation.

## 3.2 Study Design

The study design for the intervention is based on a quasi-experimental framework intended to assess changes in student learning outcomes associated with exposure to the AI-enabled adaptive learning platform. The design incorporates treatment and control groups and applies a pre- and post-intervention comparison to support attribution while remaining compatible with operational and administrative constraints in public sector school settings.

Participating schools were assigned either to a treatment group, which implemented the intervention for a month, or to a control group, which continued with standard instructional practices during the study period. Both groups were assessed at two points in time using the same standardised assessment instruments: a baseline assessment administered before the start of the intervention and an endline assessment administered following completion of the implementation period. This structure allows for the measurement of changes in learning outcomes over time within each group and the comparison of those changes between groups.

The primary analytical approach underpinning the study design is DiD framework. This approach estimates the intervention effect by comparing the change in outcomes over time in treatment schools relative to the change observed in control schools. By focusing on differences in changes rather than absolute levels, the design seeks to account for baseline differences between groups and for time-related factors that may affect all schools simultaneously.

## 3.3 Sampling Strategy

The sampling strategy was designed to ensure comparability between treatment and control groups while remaining operationally feasible and aligned with the intervention delivery model. Given that the intervention relied on a web-based AI learning platform and computer-based assessments, the **school** was adopted as the primary unit of assignment.

The study sample was restricted to **Sahiwal district**, the designated intervention district for the pilot. The sampling frame included only public high schools with **functional computer laboratories**, as these schools possessed the minimum infrastructure required to administer digital assessments and facilitate platform-based learning activities.

From the universe of eligible high schools in Sahiwal, 100 schools were **randomly selected** and subsequently **randomly assigned** to either treatment or control groups with each group having 50 schools. Random assignment at the school level was intended to minimise selection bias and to support credible comparison between intervention and non-intervention schools within the quasi-experimental design.

**Figure 1** presents the sampling approach adopted for the study.

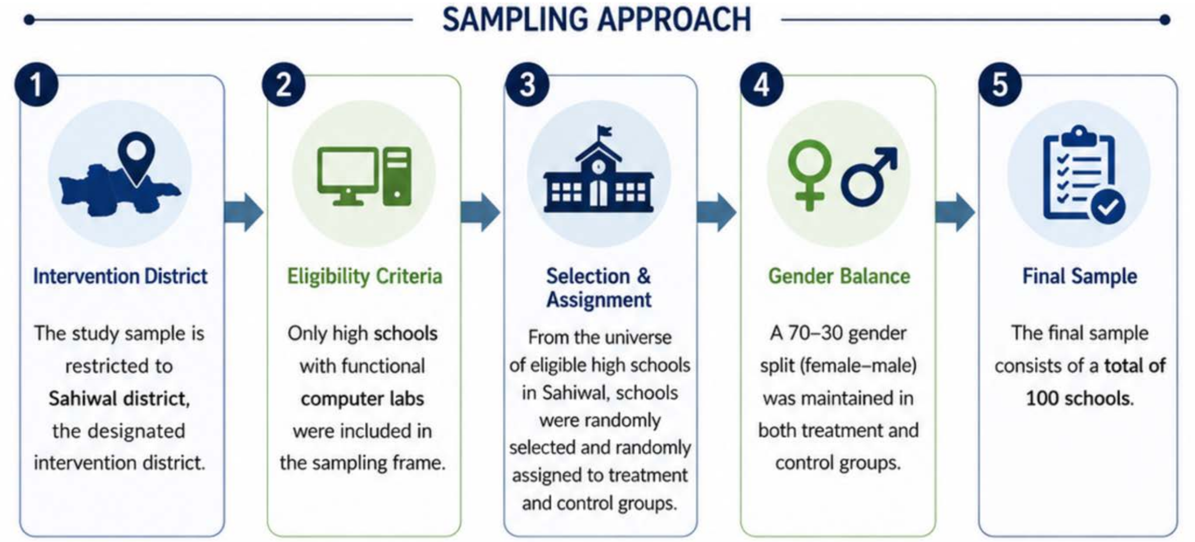



*Figure 1: Sampling Approach*

To maintain balance between the study arms and reflect the gender composition of eligible schools in the district, a **70:30 female-to-male school ratio** was maintained in both treatment and control groups.

The final sample comprised **100 public high schools**, evenly divided between treatment and control groups. Of these:

- **50 schools** were assigned to the treatment group;
- **50 schools** were assigned to the control group;
- Each group comprised **35 female schools (70 %)** and **15 male schools (30 %)**.

The composition of the final sample is presented in **Figure 2**.

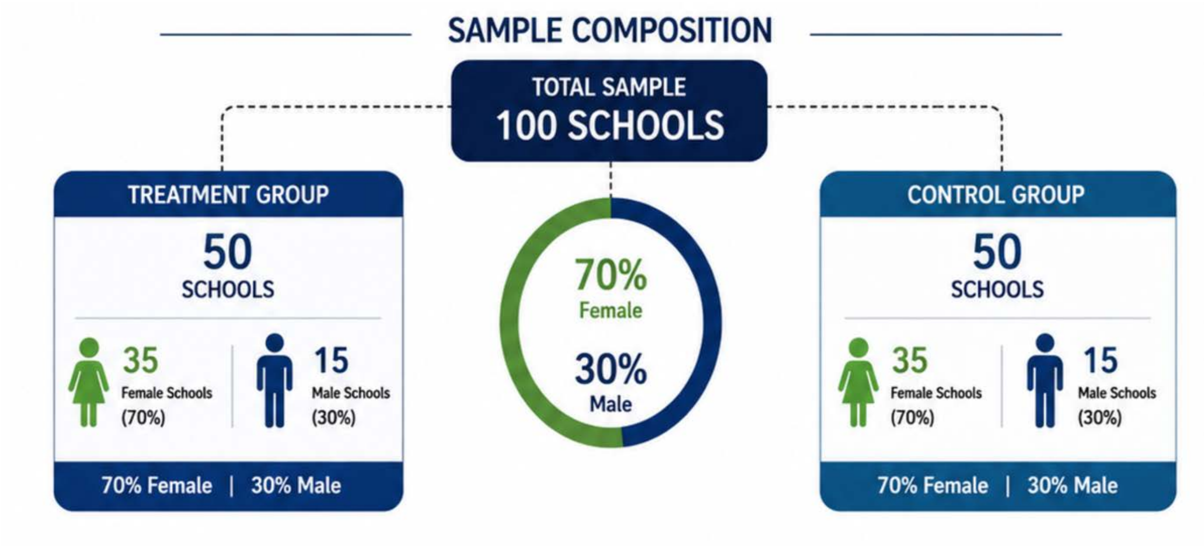



*Figure 2: Sample Composition*

## 3.4 Data Sources

The primary quantitative data source comprises standardised student learning assessments administered at baseline and endline. Assessment instruments were developed for Grade 8 English, Mathematics, and Science and aligned with approved curricula and SLOs. The same instruments were used at both time points to ensure comparability over time. Assessments were administered through the AI-enabled platform under standardised protocols, with each student accessing the assessment using a unique identifier.

Platform-generated administrative data constitute a second key data source. These data include records of student logins, assessment completion, interaction with learning content, and system usage patterns during the implementation period. Platform data were used to document exposure to the intervention and to support monitoring of implementation fidelity, rather than to generate independent outcome measures.

Qualitative inputs were used to capture perceptions of platform usability, assessment administration processes, and practical challenges encountered during delivery.

# 4. Implementation Design and Processes

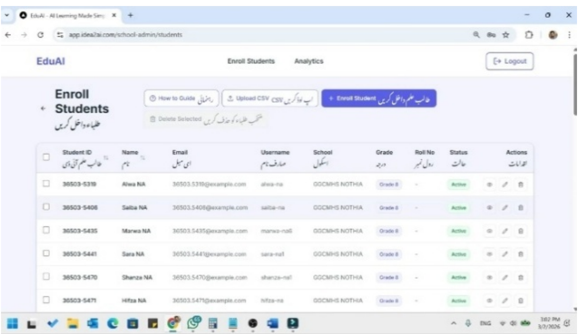
## 4.1 Adaptation of AI Enabled Learning Platform

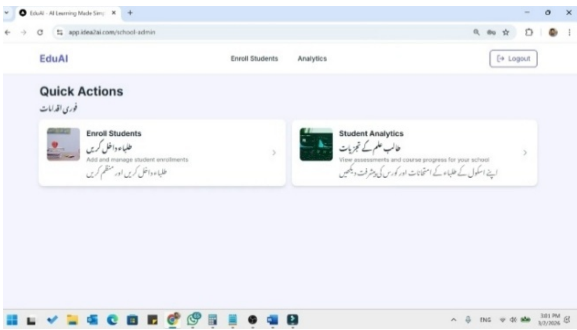
B-TAG supported the implementation of an AI-enabled learning platform for Grade 8 students in Punjab. The platform was customised using curriculum-aligned instructional content for English, Mathematics, and Science across Grades 6–8 to provide targeted remedial support based on individual student learning needs. Using students' baseline assessment results, the platform generated adaptive learning pathways tailored to each learner and continuously tracked progress throughout the intervention.

The platform incorporated several features to enhance accessibility and student engagement, including bilingual support in English and Urdu, voice-enabled navigation, and a user-friendly interface. Students could access the platform using unique login credentials, enabling personalised learning experiences and continuous monitoring of their progress. The platform could be accessed both through school computer laboratories and remotely using mobile devices, allowing students to continue learning beyond the classroom where feasible.

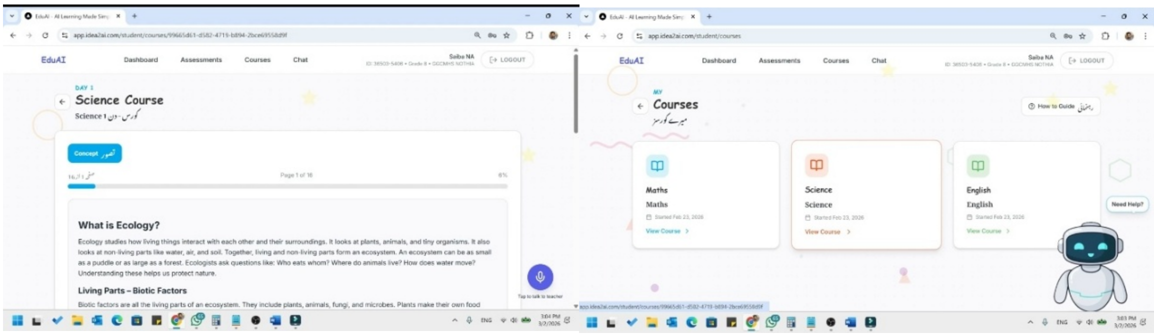
As a web-based application, the platform required no local software installation and could be accessed through commonly available internet browsers. Browser settings enabling audio input, cookies, and local storage were required to support authentication and maintain user sessions. This approach minimised technical complexity and reduced dependence on specialised IT support within schools. Reliable internet connectivity was essential for accessing the platform and its learning resources.

Separate access privileges were established for teachers and administrators responsible for student registration, platform administration, and implementation oversight. Teachers and authorised users could enrol students, monitor usage and learning progress, and access implementation reports using standard computing devices.






*Platform Screenshots – Teacher view*




*Platform Screenshots – Student view*

## 4.2 Content Development and Review

To support the effective implementation of the platform, baseline and endline assessment items were developed for prioritised SLOs across Grades 6–8 in English, Mathematics, and Science. Alongside the assessments, the platform integrated curated learning resources aligned with the identified SLOs to support personalised remediation. These resources included existing textbooks, teacher guides, and carefully selected open-source materials, all mapped to the relevant learning outcomes. Following the diagnostic assessment, the platform utilised these resources to generate structured, adaptive learning pathways tailored to individual student needs, ensuring that targeted remedial support was available across the selected grades and subjects.

A comprehensive quality assurance process was undertaken to verify the accuracy, quality, and curricular alignment of both the assessment items and learning resources. Assessment specialists from PECTAA reviewed the materials to ensure their alignment with approved SLOs, curriculum standards, grade-level appropriateness, and technical quality. Feedback from the review was systematically incorporated into subsequent revisions, ensuring that the content met quality standards and was suitable for implementation in public sector classrooms.

## 4.3 Teacher Training

Following the development of student learning content, including curated instructional resources for Grades 6–8 and baseline and endline assessment items, a series of instructional videos were developed in close collaboration with PECTAA to support implementation. The videos introduced the objectives and design of the AI-enabled learning intervention, demonstrated platform navigation for teachers and headteachers, and provided guidance on student registration, platform usage, assessment administration, and classroom facilitation.

A blended training programme was designed to ensure that participating teachers were equipped to implement the intervention effectively. Training covered platform navigation, student registration and login procedures, administration of baseline and endline assessments, facilitation of scheduled computer lab sessions, and monitoring of student progress. Teachers were also oriented to the purpose of the intervention, the appropriate use of AI-enabled learning within the school timetable, and the distinction between platform-supported learning activities and routine classroom instruction.

The training was delivered through a blended learning approach, combining self-paced instructional videos with face-to-face facilitation by the B-TAG and PECTAA team. The training was delivered in two tehsils of Sahiwal, supported by the PECTAA district field office. 78 IT teachers (26 male and 52 female) from the 50 participating schools in Sahiwal district were trained through a 1-day blended training to support the implementation of the AI-enabled adaptive learning platform.

## 4.4 Conduct of Pre Assessment

Pre-assessments were conducted in January 2026 in both treatment and control schools to establish baseline measures of student learning in English, Mathematics, and Science and to generate diagnostic information to inform subsequent platform-based learning pathways.

A total of 11 facilitators (8 male, 3 female) were engaged in Sahiwal to facilitate pre- and post-assessment conducted through computer systems in schools using each student's unique user ID. Facilitators were trained on the use of standardised protocols and monitoring tools to ensure assessment integrity, appropriate testing conditions and readiness of required hardware and software. They facilitated the assessment sessions, supporting students with basic navigation where required while adhering to guidelines intended to preserve assessment integrity. Where necessary, assessments were conducted in multiple sessions to accommodate shared device access and to ensure that all participating students had an opportunity to participate.

## 4.5 Platform-Based Learning Facilitation

Following the completion of baseline assessments, students in treatment schools engaged with the AI-enabled adaptive learning platform during scheduled learning sessions facilitated by teachers, primarily in school computer laboratories. Using their unique login credentials, students also accessed the platform remotely through mobile phones and laptops, enabling continued learning beyond the classroom where feasible.

Based on each student's baseline assessment results, the platform generated personalised learning pathways and delivered subject- and topic-specific learning resources in English, Mathematics, and Science. Learning content was sequenced according to individual learning needs and aligned with relevant SLOs, ensuring that students received targeted remedial support. Student interaction with the platform was primarily text-based, with voice-enabled features available where technical conditions permitted.

Teachers and designated school staff supported the implementation by facilitating student access to the platform, monitoring participation, and assisting students as required during learning sessions. They were able to monitor individual student progress and engagement with the platform through their unique login credentials, enabling them to provide timely support where needed.

In schools with a limited number of functional computers, students accessed the platform through shared computers using staggered schedules or small-group sessions, ensuring that all participating students had opportunities to engage with the learning activities.

Throughout the implementation period, the platform captured data on student participation, learning progress, engagement with instructional content, and completion of assigned activities. These data supported routine implementation monitoring and provided valuable evidence on student learning patterns and platform utilisation.

*Figure 3: Students using the AI-enabled learning platform in a school computer lab in Sahiwal*

The District PECTAA Office and the CEO (Education) Office in Sahiwal regularly monitored the implementation of the intervention to ensure its smooth execution and to address operational issues as they arose. The B-TAG team also conducted periodic monitoring visits to participating schools, providing technical and implementation support where required.

Monitoring visits indicated a high level of engagement among both teachers and students with the AI-enabled learning platform. While implementation was generally successful, schools reported challenges related to the limited availability of functional computers and inconsistent internet connectivity. Despite these constraints, many students reported accessing the platform from home using their parents' mobile phones, with several indicating that they used the platform for at least one hour each day beyond school hours.

**Grade 8 Female Student:** *"The best thing about this platform is that I can ask it anything without fear of being judged or feeling hesitant. Even when my answer is incorrect, I don't feel afraid of making mistakes. I had been struggling to understand terminal and non-terminal decimals even after learning them in class, but the platform explained the concept in a way that helped me understand it. Now I feel confident about it."*

**Grade 8 Female Student:** *"I enjoy using the chatbot's voice feature and often interact with it in English. Those conversations help me understand concepts much more* *clearly. The practice questions on the platform also helped me prepare for my Grade 8 board examinations, and I performed well. I would love to have a similar platform available when I move to Grade 9."*

## 4.6 Conduct of Post-Assessment

Post-assessment was conducted in March 2026 in both treatment and control schools using the same standardised assessment instruments. Administration protocols applied during the pre-assessment phase were retained to support comparability and minimise procedural variation. Assessment conduct was facilitated by the trained facilitators. Where necessary, assessments were conducted in multiple sessions to accommodate shared device access.

Completion of post-assessments concluded the field implementation phase of the intervention. Endline data generated through this process form the basis for subsequent analysis under the approved evaluation framework, which is reported in section 6.

# 5. Results and Analysis

The baseline and endline assessments administered to Grade 8 students in English, Mathematics, and Science were compared to evaluate the impact of the intervention. The results are presented progressively, beginning with overall trends in student performance, followed by comparisons between treatment and control groups, estimates of the intervention effect using a DiD approach, and disaggregated analysis by gender. Formal DiD regression results are also presented to assess the statistical significance of observed changes.

## 5.1 Overall Performance Across Subjects

Across all students, average scores improved between baseline and endline in each of the three subjects. English scores increased from **21.4 to 23.7 marks out of 50**, representing an improvement of **2.3 marks**. Mathematics scores increased from **22.9 to 25.3 marks**, an improvement of **2.4 marks**, while Science scores increased from **23.7 to 25.9 marks**, corresponding to a gain of **2.2 marks**.

*Figure 4: Scores (out of 50) by Subject – Overall*

Overall, students demonstrated improvements in all three subjects over the intervention period, with the highest endline average observed in Science (25.9 marks) and the lowest in English (23.7 marks).

## 6.2 Performance by Treatment and Control Groups

Disaggregating the results by study group shows improvements in both treatment and control schools; however, the magnitude of improvement was consistently greater among students exposed to the AI-enabled adaptive learning intervention.

- In **English**, students in control schools improved from 21.5 to 23.1 marks, while students in treatment schools improved from 21.4 to 24.3 marks

- In **Mathematics**, control group scores increased from 22.9 to 24.1 marks, compared with an increase from 23.0 to 26.4 marks in treatment schools
- In **Science**, scores in control schools increased from 23.9 to 25.2 marks, whereas treatment schools improved from 23.6 to 26.6 marks

*Figure 5: Scores (out of 50) by Subject and Group*

Improvements in both groups were expected, as students continued receiving regular classroom instruction throughout the study period. However, the consistently larger gains observed in treatment schools suggest that the AI-enabled adaptive learning intervention provided additional learning benefits beyond those achieved through routine teaching alone.

## 6.3 Absolute Changes in Scores

The change in scores between baseline and endline further highlights differences between treatment and control schools.

In the control group, average scores increased by:

- **1.60 marks** in English;
- **1.22 marks** in Mathematics; and
- **1.33 marks** in Science.

In comparison, students in the treatment group improved by:

- **2.86 marks** in English;
- **3.39 marks** in Mathematics; and
- **3.04 marks** in Science.

*Figure 6: Change in Scores (out of 50) over Time by Subject and Group*

The largest gain in the treatment group was observed in Mathematics, where students improved by more than three marks over the implementation period. This indicates that adaptive learning is particularly well suited to mathematics, where concepts are learned sequentially and immediate feedback helps students identify and address specific learning gaps before progressing to more advanced topics. Similar, although slightly smaller, improvements in Science and English indicate that the intervention was beneficial across all subjects.

## 6.4 Difference-in-Differences Estimates

The Difference-in-Differences approach estimates the additional improvement attributable to the intervention after accounting for changes observed in control schools.

The DiD estimates indicate that participation in the intervention was associated with additional gains of:

- **1.26 marks** in English;
- **2.17 marks** in Mathematics; and
- **1.71 marks** in Science.

*Figure 7: Change in Scores for Treatment Relative to Control (DiD)*

The largest intervention effect was observed in Mathematics, where treatment group students outperformed the expected improvement based on control group trends by more than two marks.

## 6.5 Percentage Changes in Scores

Expressed in percentage terms, control schools experienced relatively modest improvements across all subjects:

- **7.5 %** in English;
- **5.3 %** in Mathematics; and
- **5.6 %** in Science.

By comparison, treatment schools demonstrated substantially larger improvements:

- **13.4 %** in English;
- **14.7 %** in Mathematics; and
- **12.9 %** in Science.

*Figure 8: Percentage Change in Scores over Time by Subject and Group*

These findings indicate that percentage improvements in treatment schools were approximately double those observed in control schools across all three subjects.

The additional percentage improvement attributable to the intervention, as estimated through the DiD framework, amounted to:

- **5.9 %** in English;
- **9.4 %** in Mathematics; and
- **7.3 %** in Science.

*Figure 9: Percentage Change in Scores for Treatment Relative to Control (DiD)*

Again, Mathematics recorded the largest relative improvement attributable to the intervention.

The consistency of the findings across both absolute and percentage measures provides greater confidence in the observed pattern of results. The intervention produced positive learning gains in all three subjects, with the strongest effects in Mathematics and substantial improvements in Science. Although English also improved significantly, the relatively smaller intervention effect may reflect the more complex nature of language development, which is influenced by factors such as reading habits, vocabulary exposure, and opportunities to use language beyond the adaptive learning platform.

These findings suggest that AI-enabled adaptive learning can complement classroom instruction across multiple subjects, while offering particular value in domains that benefit from structured, sequential practice and continuous formative feedback.

## 6.6 Subject-Specific Trends

The subject-specific trend lines represent the changes in performance over time and divergence between treatment and control groups.

### *6.6.1 English*

At baseline, the two groups were virtually identical, with average scores of **21.5 marks** in the control group and **21.4 marks** in the treatment group. By endline, scores had increased to **23.1 marks** and **24.3 marks**, respectively.

Figure 10: English Scores by Group over Time

The steeper trajectory observed in the treatment group indicates that students exposed to the intervention experienced greater improvements than those in control schools. The similar baseline scores strengthen confidence that the greater improvement observed in the treatment group is associated with the intervention and adaptive learning contributed positively to language learning.

### 6.6.2 Mathematics

Baseline performance was also comparable in Mathematics, with average scores of **22.9 marks in the control group** and **23.0 marks in the treatment group**. By endline, control group scores had increased to **24.1 marks**, while treatment group scores reached **26.4 marks**.

Figure 11: Mathematics Scores by Group over Time

The widening gap between the two groups over time suggests a substantial intervention effect in Mathematics. This pattern is consistent with the sequential nature of mathematical learning, where personalised practice, immediate feedback, and targeted revision of prerequisite concepts can support steady progression and reduce persistent misconceptions.

### *6.6.3 Science*

In Science, baseline scores were 23.9 marks in control schools and 23.6 marks in treatment schools. At endline, average scores increased to 25.2 marks and 26.6 marks, respectively.

*Figure 12: Science Scores by Group over Time*

As with the other subjects, the treatment group exhibited a larger increase in scores over the implementation period suggesting that personalised learning pathways helped students reinforce scientific concepts and address individual learning gaps.

## 6.7 Gender-Based Performance

### *6.7.1 Female Students - Overall*

Among female students, average scores increased across all three subjects:

- English scores increased from **21.9 to 23.2 marks**,
- Mathematics scores increased from **22.4 to 24.5 marks**, and
- Science scores improved from **23.9 to 25.5 marks**

*Figure 13: Scores (out of 50) by Subject for Female Students*

The largest improvement among female students was observed in Mathematics, with an increase of 2.1 marks.

### *6.7.2 Male Students - Overall*

Male students recorded higher average scores than female students at both baseline and endline.

- English scores increased from **20.4 to 25.0 marks**,
- Mathematics scores increased from **24.2 to 27.6 marks**, and
- Science scores increased from **23.3 to 27.2 marks**

*Figure 14: Scores (out of 50) by Subject for Male Students*

The largest improvement among male students was observed in English, where scores increased by 4.6 marks over the intervention period.

To further explore patterns of learning improvement, assessment results for students in the treatment group were disaggregated by gender. The findings indicate that both female and male students demonstrated improvements across all three subjects over the implementation period.

### *6.7.3 Female Students – Treatment*

Among female students in treatment schools, average scores increased across all subjects.

- **English** scores improved from 22.5 to 24.5 marks, representing an **increase of 2.0 marks**
- **Mathematics** scores increased from 23.6 to 26.2 marks, corresponding to a **gain of 2.6 marks**, while
- **Science** scores improved from 24.4 to 26.7 marks, an **increase of 2.3 marks**

*Figure 15: Scores (out of 50) by Subject for Female Students – Treatment Group*

The largest improvement among female students was observed in Mathematics, where average scores increased by 2.6 marks over the intervention period.

### *6.7.4 Male Students – Treatment*

Male students in treatment schools also recorded substantial improvements across all three subjects.

- **English** scores increased from 18.8 to 23.7 marks, representing the **largest gain of 4.9 marks**
- **Mathematics** scores improved from 21.6 to 27.1 marks, an **increase of 5.5 marks**, while
- **Science** scores increased from 21.7 to 26.5 marks, corresponding to a **gain of 4.8 marks**

*Figure 16: Scores (out of 50) by Subject for Male Students – Treatment Group*

Overall, male students exhibited larger absolute gains than female students across all three subjects, with the most pronounced improvement observed in Mathematics. This can be further seen when we estimate treatment effect separately by gender using the DiD approach. The analysis compares improvements observed among students in treatment schools with those observed among students in control schools, thereby isolating the additional gains associated with exposure to the AI-enabled adaptive learning platform. Among female students, the intervention generated additional gains of **1.4 marks in English, 1.2 marks in Mathematics,** and **1.4 marks in Science** relative to changes observed among female students in control schools. For male students, the treatment effects are more pronounced, particularly in Mathematics. Male students in treatment schools achieved an additional gain of **0.8 marks in English, 4.2 marks in Mathematics,** and **1.9 marks in Science** compared with their counterparts in control schools. The substantially larger effect in Mathematics suggests that the adaptive learning platform may have been particularly effective in supporting mathematical learning among male students.

Several factors may have contributed to this difference, including variations in baseline proficiency, engagement with the platform, access to digital devices outside school, confidence in using technology, or differences in time spent practising. Future research incorporating platform usage data and qualitative evidence can help determine the extent to which these factors influenced learning outcomes

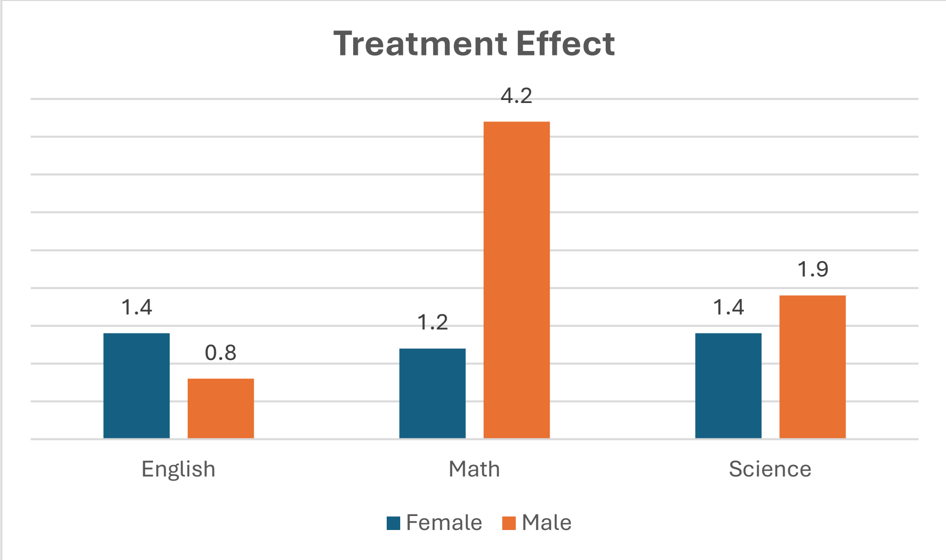



*Figure 17: Treatment effect- Comparison between boys and girls*

Notwithstanding these differences in the magnitude of gains, both female and male students demonstrated consistent improvements in learning outcomes following participation in the AI-enabled adaptive learning intervention.

## 6.8 Difference-in-Differences Regression Results

To formally assess the impact of the intervention, DiD regression models were estimated separately for each subject. The interaction term between treatment status and time captures the additional change in scores attributable to the intervention after accounting for baseline differences and general improvements over time.

The regression results indicate that the **intervention had a positive and statistically significant effect across all three subjects**.

For **English**, the estimated treatment effect was 1.259 marks ($\beta = 1.2598$, $p = 0.015$), indicating that students in treatment schools improved by an additional 1.26 marks relative to students in control schools after controlling for baseline differences.

-> SUBJECT = ENGLISH

| Source | SS | df | MS | | | |
|---|---|---|---|---|---|---|
| Model | 8636.96866 | 3 | 2878.98955 | Number of obs | = | 6,132 |
| Residual | 621826.484 | 6,128 | 101.47299 | F(3, 6128) | = | 28.37 |
| | | | | Prob > F | = | 0.0000 |
| Total | 630463.453 | 6,131 | 102.832075 | R-squared | = | 0.0137 |
| | | | | Adj R-squared | = | 0.0132 |
| | | | | Root MSE | = | 10.073 |

| Score | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| 1.treat | -.0952814 | .3443511 | -0.28 | 0.782 | -.7703304 | .5797676 |
| 1.time | 1.604269 | .3734022 | 4.30 | 0.000 | .8722692 | 2.336268 |
| treat#time | | | | | | |
| 1 1 | 1.259835 | .5185222 | 2.43 | 0.015 | .2433495 | 2.276321 |
| _cons | 21.4858 | .2476146 | 86.77 | 0.000 | 21.00039 | 21.97121 |

For **Mathematics**, the estimated treatment effect was 2.173 marks ($\beta = 2.1736$, $p < 0.001$), representing the largest intervention effect among the three subjects.

-> SUBJECT = MATH

| Source | SS | df | MS | | | |
|---|---|---|---|---|---|---|
| Model | 12615.0554 | 3 | 4205.01846 | Number of obs | = | 6,290 |
| Residual | 718384.772 | 6,286 | 114.283292 | F(3, 6286) | = | 36.79 |
| | | | | Prob > F | = | 0.0000 |
| Total | 730999.828 | 6,289 | 116.234668 | R-squared | = | 0.0173 |
| | | | | Adj R-squared | = | 0.0168 |
| | | | | Root MSE | = | 10.69 |

| Score | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| 1.treat | .1566459 | .3658921 | 0.43 | 0.669 | -.5606275 | .8739194 |
| 1.time | 1.220005 | .3913209 | 3.12 | 0.002 | .4528828 | 1.987128 |
| treat#time | | | | | | |
| 1 1 | 2.173566 | .541754 | 4.01 | 0.000 | 1.111543 | 3.235589 |
| _cons | 22.8665 | .2633375 | 86.83 | 0.000 | 22.35027 | 23.38274 |

For **Science**, the estimated treatment effect was 1.707 marks ($\beta = 1.7074$, $p = 0.001$), demonstrating a statistically significant improvement attributable to participation in the intervention.

-> SUBJECT = SCIENCE

| Source | SS | df | MS | | | |
|---|---|---|---|---|---|---|
| | | | | Number of obs | = | 6,061 |
| | | | | F(3, 6057) | = | 28.75 |
| Model | 8902.39659 | 3 | 2967.46553 | Prob > F | = | 0.0000 |
| Residual | 625267.657 | 6,057 | 103.230586 | R-squared | = | 0.0140 |
| | | | | Adj R-squared | = | 0.0135 |
| Total | 634170.054 | 6,060 | 104.648524 | Root MSE | = | 10.16 |

| Score | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| 1.treat | -.2773402 | .3518683 | -0.79 | 0.431 | -.9671272 | .4124468 |
| 1.time | 1.334069 | .3797077 | 3.51 | 0.000 | .5897073 | 2.078431 |
| treat#time | | | | | | |
| 1 1 | 1.707429 | .5253102 | 3.25 | 0.001 | .6776339 | 2.737224 |
| _cons | 23.85866 | .2535312 | 94.11 | 0.000 | 23.36164 | 24.35567 |

Overall, the findings suggest that the intervention generated meaningful improvements in learning outcomes across English, Mathematics, and Science, with the strongest effects observed in Mathematics and positive, statistically significant gains evident across all subjects.

The regression analysis confirms that the descriptive improvements observed throughout the report are statistically robust and not simply the result of overall improvements over time. The consistently positive and statistically significant treatment effects across all three subjects indicate that the AI-enabled adaptive learning intervention made an independent contribution to student learning.

# 7. Limitations

While the findings provide encouraging evidence on the effectiveness of AI-enabled adaptive learning, several limitations should be considered when interpreting the results.

- **Geographical scope**. The intervention was implemented in a single district, limiting the generalisability of the findings to other districts or provinces with different socioeconomic, infrastructural, or educational contexts. Replication across diverse settings is needed to establish external validity.

- **Short intervention duration**. The implementation period was relatively short. Although measurable learning gains were observed, a longer intervention may have produced larger effects and provided a better understanding of how learning outcomes evolve over time.

- **Selection of computer-equipped schools**. The intervention was limited to schools with functional computer laboratories and the minimum technological infrastructure required to implement the AI-enabled platform. Consequently, the findings may not fully reflect the effectiveness or feasibility of the intervention in schools with limited digital infrastructure.

- **Grade-specific evidence**. The study focused exclusively on Grade 8 students. Therefore, the results cannot be directly generalised to learners in other grades, who may have different curriculum requirements, developmental needs, and levels of digital readiness.

- **Absence of long-term follow-up**. Student learning outcomes were measured immediately after the intervention, with no follow-up assessment to determine whether the observed improvements were sustained over time. Future studies should incorporate delayed assessments to examine knowledge retention and long-term learning impacts.

- **Implementation variation**. Although implementation protocols were standardised, differences in school contexts, teacher facilitation, student attendance, technology availability, and the intensity of platform use may have influenced the magnitude of learning gains across schools. These implementation factors were not explicitly modelled in the present analysis.

- **Cost considerations**. This study focused primarily on educational effectiveness and did not include a comprehensive cost or cost-effectiveness analysis. While the intervention demonstrated positive learning gains, future research should examine implementation costs, infrastructure requirements, maintenance, teacher support, and cost-effectiveness relative to alternative instructional approaches. Such evidence would be valuable for informing decisions regarding large-scale adoption and sustainability.

# 8. Conclusions

The AI-enabled adaptive learning pilot for elementary schools demonstrates the potential of technology-enabled personalised learning to strengthen teaching and learning within Punjab's public education system. By combining curriculum-aligned diagnostic assessments, adaptive learning pathways, and real-time progress monitoring, the intervention provided students with targeted learning support while equipping teachers with actionable insights into individual learning needs. The findings indicate that students who participated in the intervention achieved significantly greater improvements in English, Mathematics, and Science than their peers in control schools, providing evidence that AI-enabled adaptive learning can contribute to improved learning outcomes when implemented within existing school structures.

Beyond its impact on student achievement, the pilot offers valuable lessons for the integration of digital learning technologies in government schools. The experience highlights the importance of teacher capacity building, reliable digital infrastructure, curriculum alignment, and continuous technical support for successful implementation. While further evidence from larger-scale implementation is needed, the pilot demonstrates that AI-enabled adaptive learning is both feasible and effective in the public sector context. Building on these findings, future efforts should focus on refining the platform, strengthening implementation mechanisms, and exploring opportunities for scaling the intervention to enable more personalised, data-driven learning across Punjab's education system.

Taken together, the implementation experience demonstrates that AI-enabled adaptive learning tools can be integrated into public sector elementary schools within bounded and clearly specified conditions. The programme’s delivery underscores the value of treating technology as a system-support mechanism, anchored in curriculum alignment, structured assessment, and institutional coordination, rather than as a standalone solution. These conclusions provide a basis for interpreting subsequent results and for informing future programme design within comparable public education contexts.

The pilot findings show that AI-enabled adaptive learning is technically feasible within Punjab's public education system, and can be implemented using existing school infrastructure, and produces measurable improvements in student learning. The findings suggest that AI should be viewed as a tool to strengthen teacher effectiveness rather than replace classroom teaching.

# 9. Recommendations

Based on the implementation experience and evaluation findings, the following recommendations are proposed to guide the institutionalisation and scale-up of AI-enabled adaptive learning within Punjab's public education system.

## 9.1 Immediate Priorities (Next 12 months)

1. **Develop a provincial roadmap for scaling AI-enabled adaptive learning.** PECTAA should develop a phased implementation strategy outlining priorities for expansion, infrastructure requirements, financing, governance arrangements, and monitoring mechanisms. The roadmap should identify priority districts and grades for future implementation based on system readiness and learning needs.

2. **Institutionalise the AI-enabled learning platform within PECTAA's digital ecosystem.** The platform should be hosted and managed through the PECTAA Data Centre (previously managed by PMIU) to ensure long-term sustainability, government ownership, data security, and integration with existing education management systems. Standard operating procedures should be established for platform maintenance, user management, technical support, and data governance.

3. **Strengthen teacher capacity to effectively use AI-enabled learning tools.** PECTAA should develop and implement structured professional development programmes that equip teachers with the knowledge and skills to integrate AI-enabled adaptive learning into classroom instruction. Training should extend beyond platform operation to include the interpretation of learner analytics, use of assessment data for instructional planning, and strategies for facilitating personalised learning.

## 9.2 Medium-Term Priorities (1-3 years)

1. **Implement a phased expansion of the intervention.** Following successful institutionalisation, PECTAA should gradually expand the AI-enabled adaptive learning model to additional districts and grade levels, prioritising schools with adequate digital infrastructure and demonstrated implementation readiness. Expansion should be accompanied by continuous monitoring, technical support, and iterative refinement based on implementation evidence.

2. **Develop integrated AI-supported teaching and learning resources.** Building on PECTAA's ongoing work on AI in education, PECTAA should develop curriculum-aligned digital teaching and learning resources that integrate AI-enabled adaptive learning with classroom instruction. These resources should include lesson plans, teacher guides, assessment activities, and differentiated learning materials linked to SLOs.

3. **Embed AI-enabled learning within teacher professional development systems.** PECTAA should integrate AI literacy, digital pedagogy, and the effective use of adaptive learning technologies into both pre-service teacher education and continuous professional development programmes. This will

ensure that teachers are equipped to effectively utilise emerging technologies as part of routine classroom practice.

## 9.3 Long-Term System Strengthening

1. **Institutionalise AI-enabled adaptive learning within Punjab's education system.** SED should position AI-enabled learning as a complementary component of the province's teaching and learning strategy by integrating adaptive learning platforms with curriculum implementation, classroom assessment, and student support systems. This should align with PECTAA's broader AI strategy and future curriculum reforms.
2. **Establish an integrated learning analytics framework.** SED and PECTAA should develop mechanisms to combine data from AI-enabled learning platforms, school-based assessments, and other education management systems to generate actionable evidence for teachers, school leaders, and policymakers. Such integration would enable more effective monitoring of student progress and inform evidence-based planning and resource allocation.
3. **Conduct long-term evaluations of effectiveness, scalability, and cost-efficiency.** Future research should assess the sustainability of learning gains, implementation fidelity at scale, and the cost-effectiveness of AI-enabled adaptive learning. Findings should inform policy decisions regarding province-wide adoption and future investments in educational technology.
4. **Strengthen partnerships for AI innovation in education.** SED should foster collaboration between government, academia, technology partners, and development organisations to continuously improve AI-enabled learning solutions, evaluate emerging technologies, and ensure that innovations remain aligned with curriculum standards, ethical considerations, and evolving educational needs.